p.Ser1235Arg should no longer be considered as a Cystic Fibrosis mutation: results from a large collaborative study

Celine Rene, Damien Paulet, Emmanuelle Girodon, Catherine Costa, Guy Lalau, Julie Leclerc, Faïza Cabet-Bey, Thierry Bienvenu, Martine Blayau, Albert Iron, et al.

p.Ser1235Arg should no longer be considered as a Cystic Fibrosis mutation: results from a large collaborative study

Céline René[1,2,3], Damien Paulet[1,2], Emmanuelle Girodon[4], Catherine Costa[4], Guy Lalau[5], Julie Leclerc[5], Faïza Cabet-Bey[6], Thierry Bienvenu[7], Martine Blayau[8], Albert Iron[29], Hervé Mittre[10], Delphine Feldmann[11], Caroline Guittard[3], Mireille Claustres[1,2,3], Marie des Georges[3]

[1] Université Montpellier1, UFR de Médecine, Montpellier, France

[2] INSERM U827, Laboratoire de Génétique de Maladies Rares, Montpellier, France

[3] CHU Montpellier, Hôpital Arnaud de Villeneuve, Laboratoire de Génétique Moléculaire, Montpellier F-34000, France

[4] Service de Biochimie-Génétique et Inserm U955 équipe 11, Groupe Henri Mondor-Albert Chenevier, APHP, Créteil, France

[5] Pôle de Biochimie et Biologie Moléculaire, Centre de Biologie Pathologie, CHU de Lille, Lille, France

[6] Service d'Endocrinologie Moléculaire et Maladies rares, Centre de Biologie et Pathologie Est, CHU de Lyon, Bron, France

[7] Laboratoire de Biochimie-Génétique, Hôpital Cochin, APHP, Paris, France

[8] Laboratoire de Génétique Moléculaire, CHU Pontchaillou, Rennes, France

[9] Laboratoire de Génétique Moléculaire, Service de Génétique Médicale, Groupe Hospitalier Pellegrin, CHU de Bordeaux, Bordeaux, France

[10] Laboratoire de Biochimie B, CHU Georges Clémenceau, Caen, France

[11] Laboratoire de Biochimie et Biologie Moléculaire, Hôpital d'Enfants Armand Trousseau, APHP, Paris, France

*Correspondence to: Céline René, Laboratoire de Génétique Moléculaire et Chromosomique, CHU, Institut Universitaire de Recherche Clinique, 641 Avenue du Doyen Gaston Giraud, 34093 Montpellier Cedex 5, France.
Telephone number: +33 4 67 41 53 60.
Fax number: + 33 4 67 41 53 65.
E-mail: celine.rene@inserm.fr

Key words: cystic fibrosis, diagnosis, genetic counseling, mutation

Running title: p.Ser1235Arg is not associated with CF disease

Abstract

Among the 1700 mutations reported in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, a missense mutation, p.Ser1235Arg is a relatively frequent finding. To clarify its clinical significance, we collected data from 104 subjects heterozygous for the mutation p.Ser1235Arg from the French CF network, addressed for various indications including classical CF, atypical phenotypes or carrier screening in subjects with or without a family history. Among them, twenty-six patients (five having CF, ten CBAVD -Congenital Absence of the Vas Deferens- and eleven with CF-like symptoms) and fourteen healthy subjects were compound heterozygous for a second *CFTR* mutation. An exhaustive *CFTR* gene analysis identified a second mutation in *cis* of p.Ser1235Arg in all CF patients and in 81.8% CBAVD patients. Moreover, epidemiological data from more than 2100 individuals found a higher frequency of p.Ser1235Arg in the general population than in CF or CBAVD patients. These data, added to the fact that *in silico* analysis and functional assays suggest a benign nature of this substitution gives several lines of evidence against an association of p.Ser1235Arg with CF or CBAVD.

Introduction

Cystic Fibrosis (CF, MIM≠219700), one of the most frequent life-shortening recessively inherited diseases, is caused by alterations in the *CFTR* gene (*Cystic Fibrosis Transmembrane Conductance Regulator*; MIM≠602421), that result in loss or impairment of CFTR-mediated ion transport across epithelial cell membrane. The clinical expression and severity of CF are highly variable. Severe phenotypes are usually associated with high concentrations of sweat chloride, early onset of pancreatic insufficiency (PI), meconium ileus at birth and severe lung disease, whereas mild phenotypes are associated with lower sweat chloride concentrations, pancreatic sufficiency (PS) and variable lung disease. Mutations in the *CFTR* gene are also involved in atypical phenotypes[1] such as CBAVD (Congenital Bilateral Absence of the Vas Deferens; MIM # 277180), mild pulmonary diseases or ICP (Idiopathic Chronic Pancreatitis).

Since the cloning of the *CFTR* gene, a wide spectrum of molecular abnormalities (more than 1700 mutations, variants or polymorphisms) has already been reported to the Cystic Fibrosis Genetic Analysis Consortium (http://www.genet.sickkids.on.ca/cftr/). However, a clear statement on the pathogenicity of a mutation is difficult to obtain, in particular for missense mutations [2,3]. Moreover, the existence of at least two mutations or sequence variations on the same allele, named complex alleles, complicates genetic counseling [4-11]. p.Ser1235Arg (3837T>G or c.3705T>G), initially reported by Cuppens *et al.* [12] with a second mutation on the same allele, p.Gly628Arg, is located in a poorly conserved region in the second nucleotide binding fold (NBF2). Earlier reports failed to establish a clear impact of p.Ser1235Arg on the phenotype severity [13-16]. The aim of this study was to determine whether we should consider p.Ser1235Arg as a possible CF-

associated mutation or reclassify it as a neutral polymorphism. Thus, we implemented a large collaborative study to repertory the patients or individuals bearing the p.Ser1235Arg mutation from nine French laboratories and compared their phenotype, genotype and associated *CFTR* haplotypes. To predict the potential effect of this sequence variation on the CFTR protein, *in silico* analyses using various programs and functional studies were performed. Here, we gather lines of evidence that p.Ser1235Arg should be no longer considered as a CF mutation.

Materials and Methods

Patients and individuals

Data were collected from one hundred and four subjects heterozygous or compound heterozygous for the p.Ser1235Arg mutation registered from the French CF network of molecular genetics laboratories. Among them, sixty-seven were referred for diagnosis: classical CF (six unrelated patients), CAVD (ten individuals with bilateral and one with unilateral Congenital Absence of the Vas Deferens), fetal suspicion of CF (eight fetuses with abnormal ultrasound signs of bowel anomalies) and forty-two individuals presenting CF-related symptoms (according to the International Classification of Disease (ICD) in the section "Cystic Fibrosis and Related Disorders" [Meeting report, 2002]) with normal or borderline sweat chloride values including genital (five individuals), respiratory (twenty-three individuals) or digestive (fourteen individuals) symptoms. Thirty-seven healthy subjects were referred for carrier screening including fourteen individuals with a positive family history and twenty-three partners of CF patients or carriers. Written consents to the genetic study were obtained from the patients and/or their family and from healthy subjects.

Molecular epidemiological study

To evaluate the p.Ser1235Arg frequency in the general population, we screened for 2114 samples: 929 anonymized dried-blood spot of neonates presenting positive or negative IRT (Immuno Reactive Trypsinogen) at day 3 obtained from the center for neonatal screening in Montpellier, South of France and 1185 genomic DNA from CF patient's or relative's partners from the cohort of the Southern France (Laboratory of molecular genetics of Montpellier) or from the cohort of the Northern France (Laboratory of molecular genetics of Lille).

Mutation Nomenclature

Gene variants at the protein level were named as recommended in the Human Genome Variation Society web page (http://www.hgvs.org/mutnomen/). For variations described at the nucleotide level, the A of the ATG translation start codon was numbered as +133 in accordance with the current *CFTR* gene numbering based on cDNA sequence (GenBank NM_000492.3) and on the CF mutation database. These variations were also given in parentheses following the approved nomenclature format (A of the ATG translation start codon as +1).

CFTR Genotype Analysis

Genomic DNA

Genomic DNA was prepared from peripherical blood leukocytes or from amniotic liquid according to standard protocols. *Patients:* The *CFTR* gene coding and flanking regions were analyzed by PCR amplification followed by DGGE (Denaturing Gel

Gradient Electrophoresis), SSCA (Single Strand Conformation Analysis), DHPLC (Denaturing High Performance Liquid Chromatography) or heteroduplex analysis followed by sequencing each abnormal pattern detected to characterize the variants. If p.Ser1235Arg was found alone, a screening for large *CFTR* rearrangements was performed using a semi-quantitative fluorescent multiplex PCR assay. *Patients' relatives*: A screening for the familial mutations was implemented. *Patients' partners:* In primary screening, five to twelve exons, including exons 13 and 19, were screened by laboratories.

To determine the haplotype associated with the p.Ser1235Arg allele, 6 microsatellite markers (IVS1(CA), IVS8(CA), IVS8(TG)m, IVS8T(n), IVS17b(TA), IVS17b(CA)) and one biallelic marker (p.Met470Val or 1540A>G(c.1408A>G)) were investigated.

Guthrie cards

Spots of 3 mm diameter, punched from anonymized cards, were distributed in 96-well plates and DNA was extracted using methanol extraction[17]. p.Ser1235Arg was screened for using DHPLC technology and each abnormal pattern was confirmed by sequencing analysis.

Computer-assisted analysis

To predict the potential pathogenicity of p.Ser1235Arg, we used PolyPhen® and SIFT® softwares. The PolyPhen® program (http://coot.embl.de/PolyPhen/) is based on the sequence homology and the mapping of the substitution site to known protein 3-dimensional structures. Results are given as "benign", "possibly damaging", "probably damaging", or "unknown". The SIFT® program (http://blocks.fhcrc.org/sift/) uses sequence homology to predict whether an amino acid substitution will affect protein

function and hence, potentially alters phenotype. Results are reported as "deleterious or not". Consequences on the splicing machinery were evaluated by Human Splicing Finder tool (http://www.umd.be/HSF). This program integrates all available matrices to identify ESE, ESS, ISE and ISS motifs: ESE Finder [18], RESCUE ESE [19], Silencer motifs from Sironi et al. [20], ESS decamers from Wang et al. [21] and new algorithms to calculate the consensus values of potential splice sites and search for Branch points. To gain some insight into the potential effect of the p.Ser1235Arg mutation from a structural point of view, we used the experimental structures of MJ0796 (in complex with ATP, pdb identifier 1l2t) as template for modeling as previously described by Eudes et al. [22]. The MJ0796 structure was visualized with MBT SimpleViewer software. The alignment of sequences of the two CFTR NBDs, other ABC transporters family and MJ0796 was obtained by ClustalW program and visualized using Jalview 2.3 tools.

In vitro functional studies

CFTR expression plasmids, Site-Directed Mutagenesis, minigene constructs and cells

A plasmid expressing the human *CFTR* gene was constructed by placing the full-length *CFTR* cDNA coding sequence obtained from pTG5985 plasmid, a gift from Transgene SA, in the pcDNA3(+) expression vector. The p.Ser1235Arg mutation was inserted into the pcDNA3-CFTR vector by site-directed mutagenesis using QuickChange® II Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions.

The human *CFTR* genomic region encompassing the exon 19 with the flanking intronic sequences, was amplified from a patient heterozygous for the p.Ser1235Arg using primers FOR 5'AATTCTGGAGCTCGAGCAGGCCTATACAGAGCCCATT-3' and

185 REV 5' CTCTTAATTTGCTAGCCATTTTCAAGATGGGAAATCTAAAACA-3'
186 (which introduce NheI and XhoI sites in the PCR product). After enzymatic digestion,
187 the PCR product was subcloned into the *Xho*I/*Nhe*I digested pSPL3 plasmid, kindly gift
188 by Dr. S. Tuffery-Giraud (Laboratory of Molecular Genetics, Montpellier, France). One
189 clone containing the wild-type allele and one with the mutation were retained for
190 expression experiments. All plasmid constructs were fully sequenced.
191 Beas2B, a human bronchial epithelial cell line expressing CFTR and the monkey kidney
192 fibroblast cells COS-7, which does not express endogenous CFTR, were cultured as
193 previously described [23]. Cells were seeded at a density of 200, 000 cells/2 ml of
194 medium in 6-well dishes. After 24h, cells were transfected with Polyfect transfection
195 reagent (Qiagen, France) according to the manufacturer's recommendations. All
196 experiments were repeated at least twice.
197 **RNA extraction and RT-PCR**
198 Total RNA was extracted using RNeasy Plus kit (Qiagen, France). Reverse transcription
199 was performed with 1 µg of total RNA as previously described [23]. 1 µl of the cDNA
200 synthesis reaction was used for PCR analyses using a mixture of 10 pmol of both
201 SD6/SA2 primers and β-actin primers, 0.5 U of Taq DNA polymerase (New England
202 Bio Labs Inc) and 10 mM of dNTPs in a 25 µl volume. Amplification was performed
203 with an initial denaturation step at 94°C for 5 min, 35 cycles of 94°C 30s, 57°C 30s,
204 72°C 90s, followed by a final extension step 72°C 7 min. PCR products were visualized
205 on 1.5 % agarose gel. Each PCR product was purified and fully sequenced.
206 **Whole cells extracts and Western-Blot**
207 Whole cells extracts from 6-well dishes were extracted and resuspended in Laemmli
208 buffer. Proteins were separated by SDS-PAGE in a 5% acrylamide gel, transferred in a

PVDF membrane and immunoblotted with antibody against CFTR (MM13-4 from Millipore).

Statistical analysis

Statistical analyses were performed using Instat® software. Chi-square test was used to compare differences between general population vs CF patients or general population vs CBAVD patients. P-values < 0.05 were considered to indicate statistical significance.

Results

Genotype/phenotype description of individuals carrying p.Ser1235Arg

Among 104 subjects from 91 unrelated families reported by the nine French laboratories, 6 patients are affected with classical CF (Table 1A). Five of them were referred for severe disease with diagnosis before one year, positive sweat test, pancreatic insufficiency and persistent respiratory infections, and one had a mild phenotype with a borderline sweat test. Complete scanning of the 27 exons and their boundaries identified a second mutation on the same allele for all of them: 1) p.Arg785X, located in exon 13, predicting a truncated protein; 2) 875+1G>A (c.743+1G>A) affecting the highly conserved splice donor site sequence; 3) the (TG)13(T)5 allele which promotes CFTR exon 9 skipping in humans [24-26].

Ten patients were referred for CBAVD (Congenital Bilateral Absence of Vas Deferens) and one for CUAVD (Congenital Unilateral Absence of Vas Deferens) without renal anomalies. The (TG)13(T)5 allele was found in 9 subjects with p.Ser1235Arg (Table 1A). Of these, eight were compound heterozygous for

232 p.Phe508del, and two for mutations associated with CF-related diseases[3,27],
233 [p.Arg117His;T7] and p.Arg1070Trp (http://www.genet.sickkids.on.ca/cftr/).

234 Among 42 patients with monosymptomatic disease carrying p.Ser1235Arg (Table
235 1A), eight cases (19%) were compound heterozygotes, seven for a severe mutation and
236 one for the intronic variant 406-6T>C (c.274-6T>C). Of these, five harbored the
237 (TG)13(T)5 allele associated in *cis* with p.Ser1235Arg. Thirty-four individuals (81%)
238 were heterozygous for p.Ser1235Arg: 5 with reduced sperm quality, 16 with
239 sinopulmonary problems such as bronchiectasis, asthma, nasal polyposis, chronic
240 obstructive pulmonary disease, and 13 with digestive symptoms (meconium ileus at
241 birth, pancreatitis, liver disease). One of them, referred for bronchitis and recurrent
242 infections without bronchiectasis, had a positive sweat test (102 mmol/l).

243 Eight fetal samples were referred for risk calculation in the cohort of
244 hyperechogenic fetal bowel. Among them, five carried only p.Ser1235Arg and were
245 considered simple heterozygote because extensive scanning did not reveal a second CF
246 mutation [28]. Three cases were compound heterozygous for another mutation. In the first
247 case, a moderate hyperechogenicity was detected at 22 weeks' pregnancy, which
248 decreased at 31. The fetus harbored the p.Phe508del mutation on the other allele and
249 parents decided to continue pregnancy. The second carried the rare missense mutation
250 p.Val920Met and depressed values of digestive enzymes were observed at 17 weeks of
251 pregnancy. In the last case, the fetus carried the complex allele
252 [p.Gly576Ala ;p.Arg668Cys], considered as a mild mutation involved in CBAVD or
253 disseminated bronchiectasis phenotypes in adults [29-31]. Although the fetus presented
254 hyperechogenic fetal bowel, the parents were reassured regarding the risk of classical
255 CF. Unfortunately, no data were available on the outcome of these pregnancies.

Among 37 individuals tested for cascade carrier screening, 14 were compound heterozygous for p.Ser1235Arg and another CFTR mutation (Table 1B), all being relatives of patients or carriers. Of these, twelve harbored in *trans* a CF-causing mutation. No symptom of CF was notified although sweat test was not performed. Notably, cases 4 and 5 were two fertile fathers carrying a Class I mutation (severe) associated in *trans* of p.Ser1235Arg.

Frequencies of p.Ser1235Arg in the general, CF and CBAVD French populations

Thirty-two p.Ser1235Arg alleles were detected in 2114 subjects of the general population representing an allelic frequency of 0.76% (Table 2). None of them had the p.Arg785X mutation. This allele frequency was compared with those of patients collated in a French study affected with CF (8/7420 alleles) or CBAVD (7/1626 alleles), 0.11% and 0.43%, respectively [32]. By χ^2 statistical analysis, the p.Ser1235Arg allelic frequency is significantly higher in the general population than in CF patients (0.76% vs. 0.11%, $p< 0.001$) but not statistically different from the CBAVD group (0.76% vs. 0.43%, $p> 0.1$). Recently, six CF patients were re-analyzed. Among them, five with severe disease, harbored the p.Arg785X mutation, while another presenting a mild phenotype had the complex allele [p.Ser1235Arg;T5]. When the p.Ser1235Arg frequencies in CF, CBAVD and the general population are compared to frequencies of p.Phe508del (Table 2), it appears that p.Ser1235Arg is as frequent as p.Phe508del ($p=0.91$) in the general population (respectively 0.76% vs. 0.82%); however, its implication in classical CF is significantly different (0.11% vs. 67.2% ($p<0.0001$) and 2.86% ($p<0.0001$) respectively).

279

280 ***p.Ser1235Arg haplotype backgrounds***

281 Haplotypic markers and familial segregation were completed in thirty-six cases
282 (Table 3). The analysis of haplotypes based on four intragenic *CFTR* markers
283 (IVS1(CA), IVS8(CA), M470V, IVS17b(CA)) revealed that p.Ser1235Arg is always
284 present on the same haplotype 26-17-M-13. This haplotype associated with the
285 polyvariant (TG)12(T)7 represents 77.8% of cases. Only the IVS17b(TA) microsatellite
286 marker shows a large variability in this study ranging from 7 to 44 repeats. However,
287 three complete haplotypes differing only by one repeat in the IVS17b(TA) (35, 36 and
288 37) accounted for more than 63% of the cases. The p.Arg785X found in CF patients in
289 *cis* of the p.Ser1235Arg occurs on a single haplotype 26-17-12-7-M-35-13. The
290 complex allele [p.Ser1235Arg;(TG)13(T)5] characteristic of CBAVD, was also found to
291 be carried on a unique haplotype (26-17-13-5-M-33-13).

292

293 ***In silico predictions and functional impact on the* CFTR *protein***

294 We evaluated the degree of conservation of serine 1235 residue using alignment
295 of the Nucleotide Binding Domains of the human CFTR with those of ABC
296 transporters, whose the CFTR belongs (Figure 1A). The results show that the serine
297 1235 is poorly conserved in the ABC transporter family, suggesting that amino acid
298 change in this position is possible. Serine 1235 residue located after a β strand is not
299 involved in anchor or structural points [33]. The analysis of its position on the 3D
300 structure of MJ0796 (Figure 1B) previously used as template for modeling the CFTR
301 NBDs [22], points that the Serine 1235 is located on the exterior surface of the NBD2.
302 Therefore, this amino acid does not participate to the NBD1/NBD2 heterodimer

formation and to the ATP binding and hydrolysis. The limited effect of this amino acid change was confirmed using SIFT® and PolyPhen® programs. To evaluate the functional impact of the mutation on the CFTR maturation, we performed Western-blot analysis on whole cell extracts from COS-7 cells transfected with either wild-type or the p.Ser1235Arg constructs (figure 1C). The results show the presence of the fully-glycosylated CFTR protein (Band C) with the p.Ser1235Arg construct, suggesting that the maturation of the pSer1235Arg-CFTR protein is not altered. These results suggest that p.Ser1235Arg has no functional impact on the CFTR protein.

Functional impact of the pSer1235Arg on the splicing

According to the Splicing Sequences Finder tools [34], p.Ser1235Arg does not seem to impact the splicing of *CFTR* transcript. To verify this prediction, we used a minigene-based approach. Constructs containing the wild type and mutant *CFTR* exon 19 and the intron-exon boundaries were inserted in the pSPL3 vector (figure 2A). After transfection in Beas2B cell line, RNAs were extracted and subjected to RT-PCR. As shown in figure 2B, compared to the wild-type, the presence of the pSer1235Arg alteration does not modify the exon 19 inclusion in the mRNA. Together, these data suggest that pSer1235Arg has no effect on the *CFTR* mRNA splicing.

Discussion

Since its initial description in 1993, the clinical significance of the p.Ser1235Arg mutation remains unclear. This ignorance makes genetic counseling very difficult, especially in the context of prenatal diagnosis. Through genotype/phenotype analysis of patients or individuals carrying p.Ser1235Arg, analyzed by the French network and

functional analysis, we present several lines of evidence that argue against an association of pSer1235Arg with CF or CBAVD although a partial penetrance of p.Ser1235Arg could not be ruled out.

First, the presence of a complex allele [p.Ser1235Arg;class I mutation] in *trans* of a severe mutation in five patients presenting severe disease (Table 1A) strongly suggests that p.Ser1235Arg alone cannot be considered as a CF-causing mutation. Moreover, among the eleven patients referred for CAVD, nine are compound heterozygous for a mutation and the complex allele [p.Ser1235Arg;(TG)13(T)5]. The (TG)13(T)5 allele without the contribution of p.Ser1235Arg is sufficient to result in male infertility or in mild phenotypes[24,25,35]. Among the two patients with p.Ser1235Arg as single allele, one has the mutation p.Phe508del in *trans* while no other mutation was identified in the second one. The possibility of mutations in intronic regions undetectable by our methods cannot be excluded.

Second, we present 14 asymptomatic individuals (Table 1B), of whom 12 carried a severe and two a mild mutation in *trans*. Three, carrying the complex allele [p.Ser1235Arg; (TG)13(T)5], are sisters or brother of CBAVD or CUAVD patients and no respiratory or digestive symptom was reported. In females, allele (T)5 has been shown to have a low penetrance [36] while no urogenital data was available for the male because of absence of parental project. Three adult males are compound heterozygous for p.Ser1235Arg and a severe mutation; two (cases 4 and 5) are biological fathers of CF children and transmitted the severe mutation. Case 7 to 11 are patients' mothers or relatives with no evidence of CF symptoms. Cases 12 and 13 are non-symptomatic sister and first-cousin of a patient with respiratory symptoms whose CF diagnostic could be discussed because symptoms are close to asthma.

Third, the frequency of the p.Ser1235Arg in the general population (Table 2) is higher than in CF and CBAVD patients, in accordance with previous reports [14,37], and is close to the p.Phe508del frequency (0.82%). By contrast, in CF or CBAVD groups, the p.Ser1235Arg frequency is significantly lower than p.Phe508del.

In the functional point of view, results from alignment show that this residue is poorly conserved and occurs in a region not implicated in known CFTR functions. Moreover, *in silico* predictions suggest that this nucleotide change is not critical for the mRNA splicing and the CFTR function. These data are confirmed by functional studies, which showed that p.Ser1235Arg does not alter the *CFTR* mRNA correct splicing and the protein maturation. These data are consistent with previous functional study of this locus in combination with alleles found at p.Met470Val and p.Gly628Arg loci [13]. Wei *et al.* demonstrated that the p.Ser1235Arg CFTR protein does not cause change in the chloride transport activity. Besides, the p.Gly628Arg/p.Ser1235Arg mutant protein induces a significantly lower cAMP dependent chloride transport activity than the p.Gly628Arg mutant protein, showing the major importance of genetic background, particularly for missense mutations.

In a second step, we determined the haplotypes linked to p.Ser1235Arg and found the same haplotype in 72.4% of cases. Only the IVS17b(TA) marker presents a large variability although three haplotypes, deriving one from the other by addition or deletion of one dinucleotide (35, 36 or 37 repeats), represent more than 55% of the alleles. Besides, IVS1(CA) 26, a rare allele in the European populations (4.2% vs 80% for alleles 22, 23 ad 24) [38], is strictly linked to p.Ser1235Arg. In CBAVD patients, the haplotype 26-17-13-5-M-33-13 was found in 85.7% of p.Ser1235Arg alleles. This data

suggests the further change of the (TG)12(T)7 to (TG)13(T)5 on a chromosome bearing the alteration p.Ser1235Arg on a background IVS17b(TA)33. Together, this data suggests that p.Ser1235Arg could derive from a founding event and haplotypes differing by a single microsatellite could depend on slippage phenomena, as already proposed [39].

In conclusion, there are now strong lines of evidence against a severe deleterious effect of p.Ser1235Arg and its association with classical CF or CBAVD disease. However, a partial penetrance and/or the presence of other sequence alterations in non-screened regions particularly in patients with atypical or mild symptoms of cystic fibrosis [27,40,41] cannot be ruled-out. This data highlights the importance of searching for a complex allele whenever p.Ser1235Arg is identified. The absence of clinical effects in healthy individuals compound heterozygous for the p.Ser1235Arg and another severe mutation is helpful to provide adequate genetic counseling to the families with regard to its transmission in offspring.

Acknowledgements

We thank Dr C. Beroud for fruitful discussion in the software analysis and the GREPAM for efficient anonymized sample referral.

Conflict of interest.

The authors declare no conflict of interest.

References

1. Kerem E: Atypical CF and CF related diseases. *Paediatr Respir Rev* 2006; **7 Suppl 1:** S144-146.

2. Cotton RG, Scriver CR: Proof of "disease causing" mutation. *Hum Mutat* 1998; **12:** 1-3.

3. Castellani C, Cuppens H, Macek M, Jr. *et al*: Consensus on the use and interpretation of cystic fibrosis mutation analysis in clinical practice. *J Cyst Fibros* 2008; **7:** 179-196.

4. Savov A, Angelicheva D, Balassopoulou A, Jordanova A, Noussia-Arvanitakis S, Kalaydjieva L: Double mutant alleles: are they rare? *Hum Mol Genet* 1995; **4:** 1169-1171.

5. Fanen P, Clain J, Labarthe R *et al*: Structure-function analysis of a double-mutant cystic fibrosis transmembrane conductance regulator protein occurring in disorders related to cystic fibrosis. *FEBS Lett* 1999; **452:** 371-374.

6. Romey MC, Guittard C, Chazalette JP *et al*: Complex allele [-102T>A+S549R(T>G)] is associated with milder forms of cystic fibrosis than allele S549R(T>G) alone. *Hum Genet* 1999; **105:** 145-150.

7. Clain J, Fritsch J, Lehmann-Che J *et al*: Two mild cystic fibrosis-associated mutations result in severe cystic fibrosis when combined in cis and reveal a residue important for cystic fibrosis transmembrane conductance regulator processing and function. *J Biol Chem* 2001; **276:** 9045-9049.

8. Rohlfs EM, Zhou Z, Sugarman EA *et al*: The I148T CFTR allele occurs on multiple haplotypes: a complex allele is associated with cystic fibrosis. *Genet Med* 2002; **4:** 319-323.

9. Clain J, Lehmann-Che J, Girodon E *et al*: A neutral variant involved in a complex CFTR allele contributes to a severe cystic fibrosis phenotype. *Hum Genet* 2005; **116:** 454-460.

10. Niel F, Legendre M, Bienvenu T *et al*: A new large CFTR rearrangement illustrates the importance of searching for complex alleles. *Hum Mutat* 2006; **27:** 716-717.

11. Claustres M, Altieri JP, Guittard C, Templin C, Chevalier-Porst F, Des Georges M: Are p.I148T, p.R74W and p.D1270N cystic fibrosis causing mutations? *BMC Med Genet* 2004; **5:** 19.

12. Cuppens H, Marynen P, De Boeck C, Cassiman JJ: Detection of 98.5% of the mutations in 200 Belgian cystic fibrosis alleles by reverse dot-blot and

sequencing of the complete coding region and exon/intron junctions of the CFTR gene. *Genomics* 1993; **18:** 693-697.

13. Wei L, Vankeerberghen A, Jaspers M, Cassiman J, Nilius B, Cuppens H: Suppressive interactions between mutations located in the two nucleotide binding domains of CFTR. *FEBS Lett* 2000; **473:** 149-153.

14. Monaghan KG, Feldman GL, Barbarotto GM, Manji S, Desai TK, Snow K: Frequency and clinical significance of the S1235R mutation in the cystic fibrosis transmembrane conductance regulator gene: results from a collaborative study. *Am J Med Genet* 2000; **95:** 361-365.

15. Dequeker E, Stuhrmann M, Morris MA *et al*: Best practice guidelines for molecular genetic diagnosis of cystic fibrosis and CFTR-related disorders--updated European recommendations. *Eur J Hum Genet* 2009; **17:** 51-65.

16. Oca F, Dreux S, Gerard B *et al*: Amniotic fluid digestive enzyme analysis is useful for identifying CFTR gene mutations of unclear significance. *Clin Chem* 2009; **55:** 2214-2217.

17. Caggana M, Conroy JM, Pass KA: Rapid, efficient method for multiplex amplification from filter paper. *Hum Mutat* 1998; **11:** 404-409.

18. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR: ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res* 2003; **31:** 3568-3571.

19. Fairbrother WG, Yeh RF, Sharp PA, Burge CB: Predictive identification of exonic splicing enhancers in human genes. *Science* 2002; **297:** 1007-1013.

20. Sironi M, Menozzi G, Riva L *et al*: Silencer elements as possible inhibitors of pseudoexon splicing. *Nucleic Acids Res* 2004; **32:** 1783-1791.

21. Wang Z, Rolish ME, Yeo G, Tung V, Mawson M, Burge CB: Systematic identification and analysis of exonic splicing silencers. *Cell* 2004; **119:** 831-845.

22. Eudes R, Lehn P, Ferec C, Mornon JP, Callebaut I: Nucleotide binding domains of human CFTR: a structural classification of critical residues and disease-causing mutations. *Cell Mol Life Sci* 2005; **62:** 2112-2123.

23. Rene C, Taulan M, Iral F *et al*: Binding of serum response factor to cystic fibrosis transmembrane conductance regulator CArG-like elements, as a new potential CFTR transcriptional regulation pathway. *Nucleic Acids Res* 2005; **33:** 5271-5290.

24. Disset A, Michot C, Harris A, Buratti E, Claustres M, Tuffery-Giraud S: A T3 allele in the CFTR gene exacerbates exon 9 skipping in vas deferens and

epididymal cell lines and is associated with Congenital Bilateral Absence of Vas Deferens (CBAVD). *Hum Mutat* 2005; **25:** 72-81.

25. Groman JD, Hefferon TW, Casals T *et al*: Variation in a repeat sequence determines whether a common variant of the cystic fibrosis transmembrane conductance regulator gene is pathogenic or benign. *Am J Hum Genet* 2004; **74:** 176-179.

26. Claustres M: Molecular pathology of the CFTR locus in male infertility. *Reprod Biomed Online* 2005; **10:** 14-41.

27. Feldmann D, Couderc R, Audrezet MP *et al*: CFTR genotypes in patients with normal or borderline sweat chloride levels. *Hum Mutat* 2003; **22:** 340.

28. Muller F, Simon-Bouy B, Girodon E, Monnier N, Malinge MC, Serre JL: Predicting the risk of cystic fibrosis with abnormal ultrasound signs of fetal bowel: results of a French molecular collaborative study based on 641 prospective cases. *Am J Med Genet* 2002; **110:** 109-115.

29. Bombieri C, Benetazzo M, Saccomani A *et al*: Complete mutational screening of the CFTR gene in 120 patients with pulmonary disease. *Hum Genet* 1998; **103:** 718-722.

30. Casals T, De-Gracia J, Gallego M *et al*: Bronchiectasis in adult patients: an expression of heterozygosity for CFTR gene mutations? *Clin Genet* 2004; **65:** 490-495.

31. Pallares-Ruiz N, Carles S, Des Georges M *et al*: Complete mutational screening of the cystic fibrosis transmembrane conductance regulator gene: cystic fibrosis mutations are not involved in healthy men with reduced sperm quality. *Hum Reprod* 1999; **14:** 3035-3040.

32. Claustres M, Guittard C, Bozon D *et al*: Spectrum of CFTR mutations in cystic fibrosis and in congenital absence of the vas deferens in France. *Hum Mutat* 2000; **16:** 143-156.

33. Callebaut I, Eudes R, Mornon JP, Lehn P: Nucleotide-binding domains of human cystic fibrosis transmembrane conductance regulator: detailed sequence analysis and three-dimensional modeling of the heterodimer. *Cell Mol Life Sci* 2004; **61:** 230-242.

34. Desmet F, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C: Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009; **37:** e67.

35. Cuppens H, Lin W, Jaspers M *et al*: Polyvariant mutant cystic fibrosis transmembrane conductance regulator genes. The polymorphic (Tg)m locus

explains the partial penetrance of the T5 polymorphism as a disease mutation. *J Clin Invest* 1998; **101:** 487-496.

36. Sun W, Anderson B, Redman J *et al*: CFTR 5T variant has a low penetrance in females that is partially attributable to its haplotype. *Genet Med* 2006; **8:** 339-345.

37. Modiano G, Bombieri C, Ciminelli BM *et al*: A large-scale study of the random variability of a coding sequence: a study on the CFTR gene. *Eur J Hum Genet* 2005; **13:** 184-192.

38. Mateu E, Calafell F, Bonne-Tamir B *et al*: Allele frequencies in a worldwide survey of a CA repeat in the first intron of the CFTR gene. *Hum Hered* 1999; **49:** 15-20.

39. Claustres M, Desgeorges M, Moine P, Morral N, Estivill X: CFTR haplotypic variability for normal and mutant genes in cystic fibrosis families from southern France. *Hum Genet* 1996; **98:** 336-344.

40. Castellani C, Benetazzo MG, Tamanini A, Begnini A, Mastella G, Pignatti P: Analysis of the entire coding region of the cystic fibrosis transmembrane regulator gene in neonatal hypertrypsinaemia with normal sweat test. *J Med Genet* 2001; **38:** 202-205.

41. Reboul MP, Laharie D, Amouretti M, Lacombe D, Iron A: Isolated idiopathic chronic pancreatitis associated with a compound heterozygosity for two mutations of the CFTR gene. *Gastroenterol Clin Biol* 2003; **27:** 821-824.

42. Robertson NH, Weston SL, Kelly SJ *et al*: Development and validation of a screening test for 12 common mutations of the cystic fibrosis CFTR gene. *Eur Respir J* 1998; **12:** 477-482.

Titles and legends to figures

Figure1: Structural and processing impact of p.Ser1235Arg mutation on the CFTR protein.

A: Conservation of residues close to Ser1235 (blue square) in the two NBD sequences of human CFTR and other ABC transporters. Positions always occupied by hydrophobic amino acids (V, I, L, M, F, Y, W) are shaded blue. Conserved glycine and lysine are colored in orange and in red, respectively.

B: Localization of the serine 1235 on the 3D-modelization of the NBD1/NBD2 heterodimer .

C: COS-7 cell line were transfected with either wild-type or p.Ser1235Arg CFTR constructs. 48 hours post-tranfection, whole-cell proteins were extracted and subjected to a Western-blot analysis. Experiments were repeated three times.

Figure 2: Functional impact on the *CFTR* mRNA splicing

A: Schema of the minigene constructs in pSPL3 vector. The exon 19 and exon-intron junctions were inserted in the multiple cloning site of the pSPL3 vector between the exon A and B from the β-globin by PCR-restriction.

B: RT-PCR analysis of mRNA of cell expressing Wild-Type[42] or p.Ser1235Arg(S1235R) constructs. BeaS2B cells were transfected with either empty, wild-type or p.Ser1235Arg pSPL3 vector. After 48 hours, total RNA were extracted and cDNA were synthetized. PCR was performed using SA2 and SD6 primers amplifying the region between exon A and B. On the right, the composition of each transcript was indicated. Each experiment was repeated at least three times.

597 Table 1A Genotype and phenotype of CF, CBAVD and CF-like patients carrying the

598 p.Ser1235Arg mutation.

Genotype		phenotype	No of subjects
allele 1	allele 2		
p.Ser1235Arg;p.Arg785X	p.Phe508del	severe CF	2
p.Ser1235Arg;p.Arg785X	NA [1]	severe CF	1
p.Ser1235Arg;875+1G>A (c.743+1C>A)	3629delT (c.3497delT)	severe CF	1
p.Ser1235Arg;p.Arg785X	p.Gly542X	severe CF	1
p.Ser1235Arg;(TG)13(T)5	p.Gly551Asp	mild CF	1
p.Ser1235Arg;(TG)13(T)5	p.Phe508del	CBAVD	6
p.Ser1235Arg;(TG)13(T)5	p.Arg1070Trp	CBAVD	1
p.Ser1235Arg;(TG)13(T)5	p.Arg117His; (T)7	CBAVD	1
p.Ser1235Arg	p.Phe508del	CBAVD	1
p.Ser1235Arg	-	CBAVD	1
p.Ser1235Arg;(TG)13(T)5	p.Phe508del	CUAVD	1
		Suspicion CF/mild phenotype:	
p.Ser1235Arg	-	genital symptoms	5
p.Ser1235Arg	-	respiratory symptoms	16
p.Ser1235Arg;(TG)13(T)5	p.Phe508del	respiratory symptoms	2
p.Ser1235Arg	406-6T>C (c.274-6T>C)	respiratory symptoms	1
p.Ser1235Arg	p.Tyr1092X	respiratory symptoms	1
p.Ser1235Arg	p.Glu831X	respiratory symptoms	1
p.Ser1235Arg	p.Gln493X	respiratory symptoms	1
p.Ser1235Arg	p.Ile507del	respiratory symptoms	1
p.Ser1235Arg	-	digestive symptoms	13
p.Ser1235Arg	p.Gly542X	digestive symptoms	1
p.Ser1235Arg	-	hyperechogenic fetal bowel	5
p.Ser1235Arg	p.Arg668Cys;p.Arg576Ala	hyperechogenic fetal bowel	1
p.Ser1235Arg	p.Val920Met	hyperechogenic fetal bowel	1
p.Ser1235Arg	p.Phe508del	hyperechogenic fetal bowel	1

599 [1] NA: Not Available ; We could only test the mother and an healthy sister (the patient

600 was deceased and the father's DNA was not available)

601

602 Table 1B Genotype and familial history of asymptomatic individuals carrying the

603 p.Ser1235Arg mutation.

Case #	Genotype		Familial information
	allele 1	allele 2	
1	p.Ser1235Arg;(TG)13(T)5	p.Phe508del	brother of CUAVD (no parental project)
2	p.Ser1235Arg;(TG)13(T)5	p.Phe508del	sister of CUAVD
3	p.Ser1235Arg;(TG)13(T)5	p.Arg1070Trp	sister of CBAVD
4	p.Ser1235Arg	p.Gly542X	father of CF [p.Phe508del]+ [p.Gly542X] and healthy children [p.Ser1235Arg]+ [(TG)11(T)5]
5	p.Ser1235Arg	p.Gln493X	father of CF [p.Phe508del]+[p.Gln493X]
6	p.Ser1235Arg	p.Phe508del	uncle of CF (no parental project)
7	p.Ser1235Arg	p.Phe508del	mother of CF [p.Phe508del]+[1717-1G>A (c.1585-1G>A)]
8	p.Ser1235Arg	2347delG (c.2215delG)	mother of CF [p.Phe508del]+[2347delG]
9	p.Ser1235Arg	(TG)11(T)5	mother of non CF fetus with hyperechogenic fetal bowel [p.Phe508del]+[(TG)11(T)5]
10	p.Ser1235Arg	p.Phe508del	sister of CBAVD
11	p.Ser1235Arg	p.Phe508del	sister of CBAVD
12	p.Ser1235Arg	p.Glu831X	sister of CF-like patient
13	p.Ser1235Arg	p.Phe508del	first-cousin of CF-like patient
14	p.Ser1235Arg	p.Phe508del	18 month-old child with prenatal diagnostic based on familial history of CF

604

605

606

607

608 Table 2: Comparison of the allelic frequencies of p.Ser1235Arg and p.Phe508del in the

609 general population, CF and CBAVD patients.

	Populations			Significance (p-values)	
	General population This study	CF patients [30] n=7,420	CBAVD patients [30] n=1,626	General population vs. CF	General population vs. CBAVD
p.Phe508del	0.82% (n=1950)	67.2%	21.6%	S (p< 0.05)	S (p< 0.05)
p.Ser1235Arg	0.76 % (n=4228)	0.11%	0.43%	S (p< 0.05)	NS (p= 0.17)
Significance (p-values) p.Phe508del vs p.Ser1235Arg	NS (p=0.91)	S (p< 0.05)	S (p< 0.05)		

610 n: number of unrelated tested chromosomes. S: Significant difference; NS: No

611 Significant difference. P-values significant at <0.05

612

613

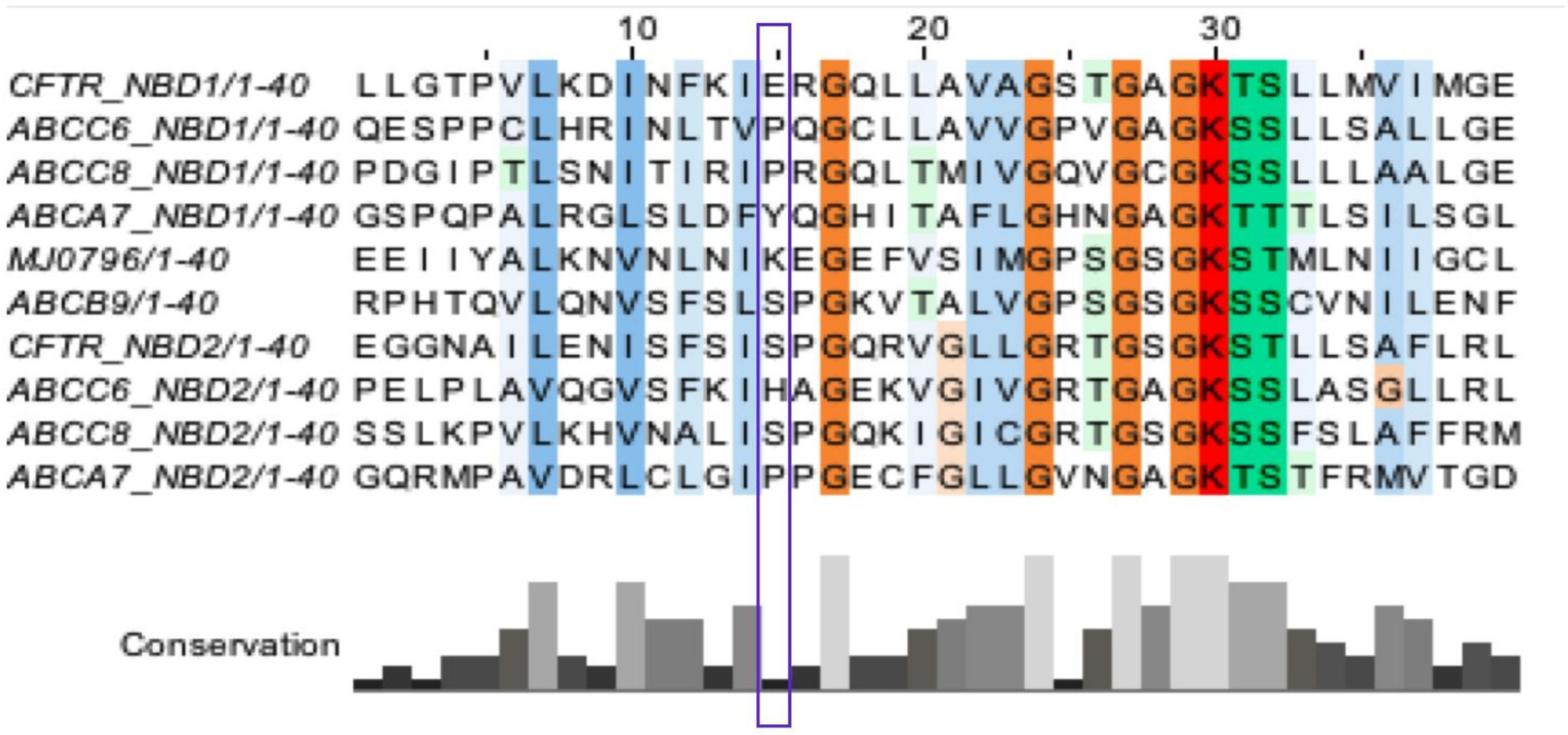
614 Table 3: *CFTR* haplotypes for 36 fully haplotyped p.Ser1235Arg chromosomes from
615 CF, CBAVD patients, ruled-out CF (including CF-like symptoms and Fetal echogenic
616 bowel) and general population.

IVS1(CA)-IVS8(CA)-M470V-IVS17b(CA)	IVS8(TG)m - IVS8(T)n	Mutation in *cis*	IVS17b(TA)	n	%
26-17-M-13	12-7	p.Arg785X	35	4	11.1
		875+1G>A (c.743+1C>A)	36	1	2.8
		None	33	1	2.8
			35	5	13.8
			36	9	25
			37	4	11.1
			38	1	2.8
			40	1	2.8
			44	1	2.8
			7	1	2.8
	13-5		33	8	22.2

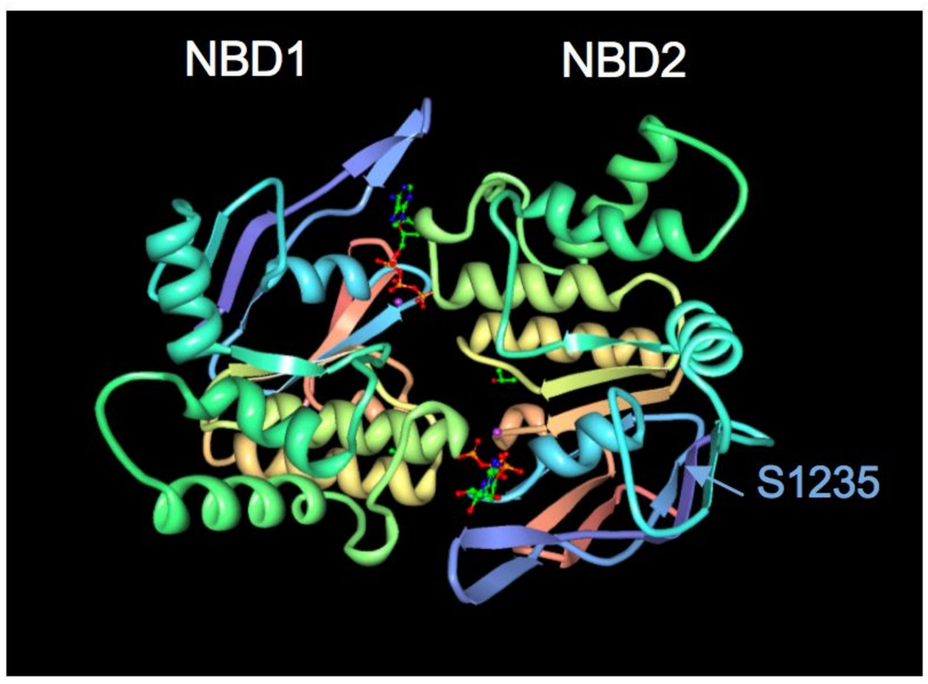
617
618
619
620
621
622

Figure 1

A


10
20
30
CFTR_NBD1/1-40 LLGTPVLKDINFKIERGQLLAVAGSTGAGKTSLLMVIMGE
ABCC6_NBD1/1-40 QESPPCLHRINLTVPQGCLLAVVGPVGAGKSSLLSALLGE
ABCC8_NBD1/1-40 PDGIPTLSNITIRIPRGQLTMIVGQVGCGKSSLLLAALGE
ABCA7_NBD1/1-40 GSPQPALRGLSLDFYQGHITAFLGHNGAGKTTTLSILSGL
MJ0796/1-40 EEIIYALKNVNLNIKEGEFVSIMGPSGSGKSTMLNIIGCL
ABCB9/1-40 RPHTQVLQNVSFSLSPGKVTALVGPSGSGKSSCVNILENF
CFTR_NBD2/1-40 EGGNAILENISFSISPGQRVGLLGRTGSGKSTLLSAFLRL
ABCC6_NBD2/1-40 PELPLAVQGVSFKIHAGEKVGIVGRTGAGKSSLASGLLRL
ABCC8_NBD2/1-40 SSLKPVLKHVNALISPGQKIGICGRTGSGKSSFSLAFFRM
ABCA7_NBD2/1-40 GQRMPAVDRLCLGIPPGECFGLLGVNGAGKTSTFRMVTGD
Conservation

B


NBD1
NBD2
S1235

C

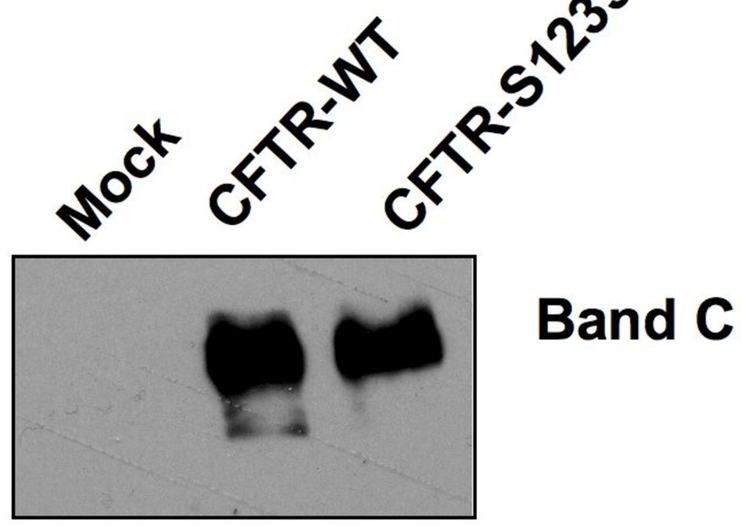
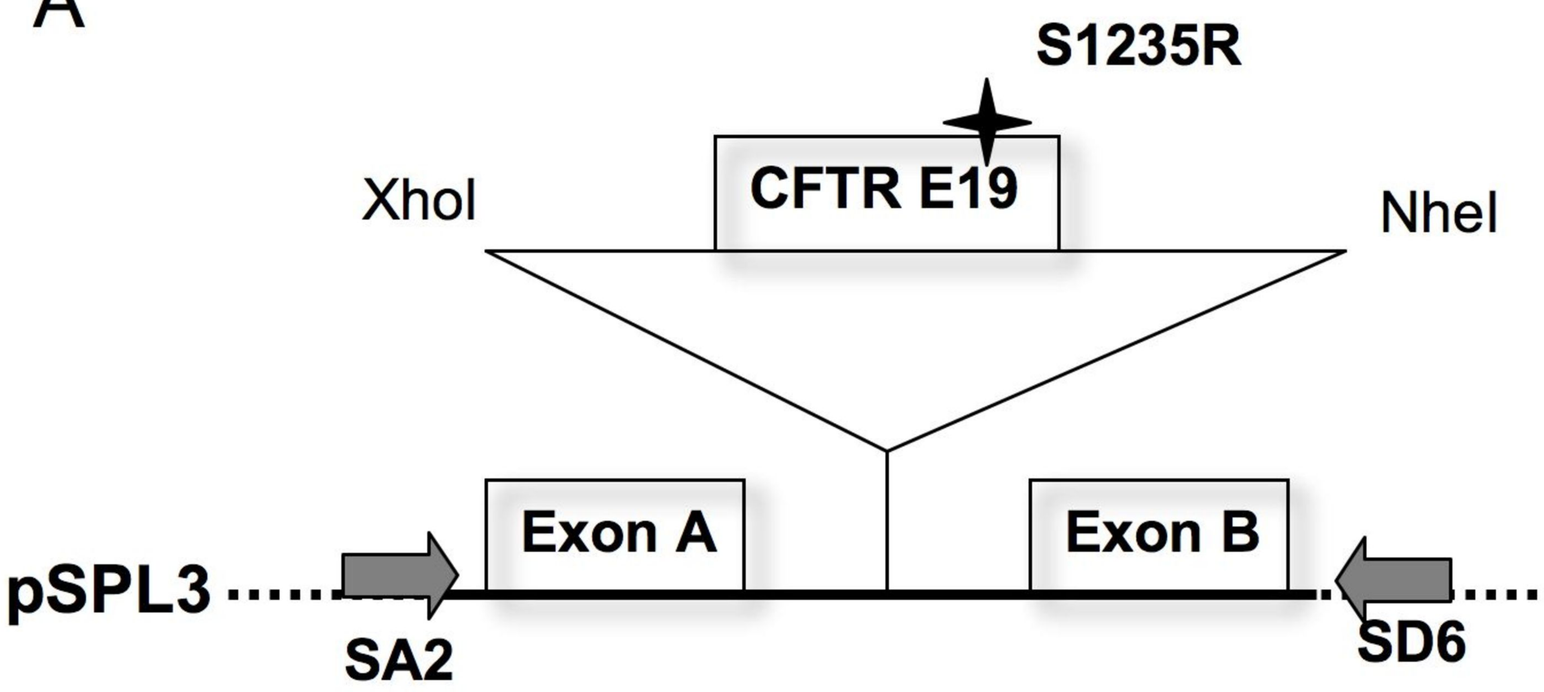
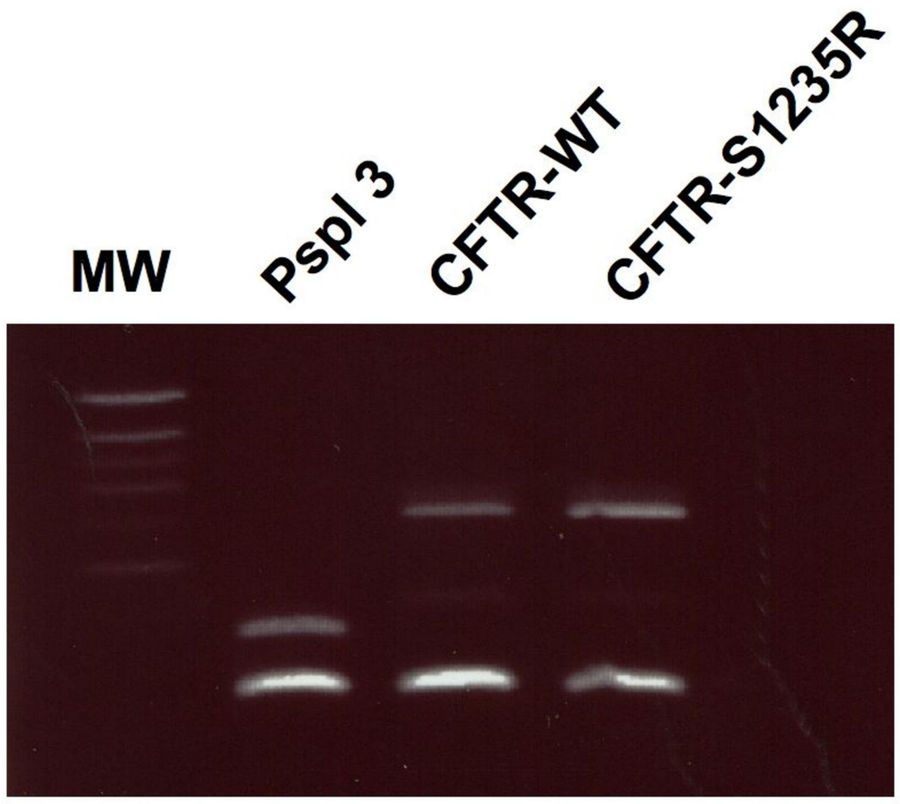

Mock
CFTR-WT
CFTR-S1235R
Band C

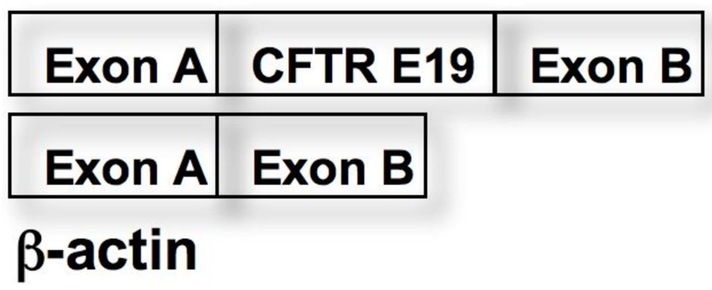
Figure 2

A


S1235R
CFTR E19
XhoI
NheI
Exon A
Exon B
pSPL3
SA2
SD6

B


MW
PspI 3
CFTR-WT
CFTR-S1235R


Exon A
CFTR E19
Exon B
Exon A
Exon B
β-actin